Surveying Immigrants and Immigration 1

Taeku Lee
taekulee@fas.harvard.edu

12 June 2023
BIMI Summer Institute
Polling and Immigration
Issue Salience – "Immigration"

Source: Gallup polls

May-96, 6%
Apr-06, 19%
Jul-06, 14%
Jun-09, 14%
Jul-2014, 17%
Jul-18, 22%
Oct-2022, 46%
Polling and Issue Salience

Source: Gallup polls and Roper iPoll database
"Should immigration be kept at its present level, increased or decreased?"

- Trend data are rare and limited.
- These trend data would suggest a period of liberalization of views on immigration during a period of deep division and heightened nativism.
Composition is not limited to surveys that target foreign-born samples or racial/ethnic groups with large immigrant 1st and 2nd sub-groups.

What may be interpreted as changes in public opinion may instead be changes in the composition of the "public"
Content: Types of Questions

- **Overall:** "out of control, or not?" "works pretty well ... or needs to be rebuilt?"
- **Levels:** "too many, too few, about the right amount?"
- **Specific policies:** "build a 2,000 mile long security fence ... to stop illegal immigration?"
- **Impact:** "good or bad for this country?" "good or bad for the economy?" "benefits from legal immigration outweigh the risks?"
- **Identity:** "immigration strengthens or weakens the American character" "Our country was founded by immigrants and we benefit from the diversity of immigration"

- Questions *about* immigration, not question *of* immigrants.
Primers

→ Surveys and "public opinion" as a concept
→ Parallel processes and dual inferences in survey research
→ Sampling
→ Measurement
Primers

→ Surveys and "public opinion" as a concept
Normative Force of Public Opinion

• **James Bryce (1895):** “Towering over Presidents and State governors, over Congress and State legislatures, over conventions and vast machinery of party, public opinion stands out, in the United States, as the great source of power, *the master of servants* who tremble before it.”

• **Alexander Hamilton (1787):** "The republican principle demands that the deliberative sense of the community guide the conduct of those to whom they entrust the management of their affairs; but it does not require an unqualified complaisance to every sudden breeze of passion, or to every transient impulse which the people may receive from the arts of men, who flatter their prejudices to betray their interest."

• **H.L. Mencken (1920):** "As democracy is perfected, the office of the President represents, more and more closely, the inner soul of the people. On some great and glorious day, the plain folks of the land will reach their heart's desire at last and the White House will be occupied by a downright fool and a complete narcissistic moron."
What is Public Opinion?

• Henry Maine (1914): “Vox Populi may be Vox Dei, but very little attention shows that there has never been agreement as to what Vox means or as to what Populus means.”

• V.O. Key (1961): “those opinions held by private persons which governments find it prudent to heed.”
## Varieties of "Opinions Held"

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oratory/rhetoric</td>
<td>5th century, BC</td>
</tr>
<tr>
<td>Printing</td>
<td>16th century</td>
</tr>
<tr>
<td>Crowds</td>
<td>17th century</td>
</tr>
<tr>
<td>Petitions</td>
<td>late 17th century</td>
</tr>
<tr>
<td>Salons &amp; coffeehouses</td>
<td>late 17th/early 18th c.</td>
</tr>
<tr>
<td>Revolutions</td>
<td>late 18th century</td>
</tr>
<tr>
<td>Strikes</td>
<td>19th century</td>
</tr>
<tr>
<td>General elections</td>
<td>19th century</td>
</tr>
<tr>
<td>Straw polls</td>
<td>1820s</td>
</tr>
<tr>
<td>Modern newspapers</td>
<td>mid-19th century</td>
</tr>
<tr>
<td>Letters to officials and editors</td>
<td>mid-19th century</td>
</tr>
<tr>
<td>Sample survey</td>
<td>1930s</td>
</tr>
<tr>
<td>Email, bulletin boards, chatrooms</td>
<td>1990s</td>
</tr>
<tr>
<td>Blogs, tweets, searches, public comments</td>
<td>21st century</td>
</tr>
</tbody>
</table>
Normative Force of Survey Research

• **Archibald Crossley (1937):** “Scientific polling makes it possible within two or three days at moderate expense for the entire nation to work hand in hand with its legislative representatives, on laws which affect our daily lives. Here is the long-sought key to “Government by the people.”

• **Sidney Verba (1996):** “Surveys produce just what democracy is supposed to produce—equal representation of all citizens.”

• **Henry Brady (2000):** "Like telescopes in astronomy, microscopes in biology, and seismic, weather, and environmental sensors in the geosciences, surveys have features that make them a fundamental data collection method for the social sciences."
"Public Opinion" in Practice

- Philip Converse (1987): “It is exactly this kind of ‘one person, one vote’ tally of opinions as routinely reported today by polls and surveys which has now become the consensual understanding of the world around as to a baseline of public opinion.”

POQ Wordle, 1937-59

Source: https://blogs.rti.org/surveypost/
POQ Wordle, 1960-89

Source: https://blogs.rti.org/surveypost/
POQ Wordle, 1990-2012

Source: https://blogs.rti.org/surveypost/
Primers

→ Surveys and "public opinion" as a concept
→ Parallel processes and dual inferences in survey research
So, You Want to Do a Survey?

Three questions to ask and answer:

→ Who do you want to ask?
→ What do you want to ask?
→ How much do you have?
Two Types of Inference

**answers to questions**

**Respondent characteristics**

**population characteristics**

**Measurement:** “How well do questions capture respondent characteristics?”

**Representation:** “How well do survey participants correspond to the target population?”
**Measurement**

**What you want to know about:**
Immigration views; experience of bias; voter turnout; consumer sentiment; etc.

**How you want to measure it:**
Questions ("Should immigration be kept at its present level, increased or decreased?")

**Which categories of reply:**
"increased / decreased / kept same"; “don't know”; refused to answer; Likert scale; feeling thermometer; factor scores / latent variables; list experiments.

**What you actually analyze:**
After outliers, coding errors, processing errors, etc., these are final data from which inference is drawn about the construct for an individual respondent.
Representativeness

**Who** you want to study (e.g., adult citizens in the US; immigrant 1.5 generation Latinx in Texas; high-skilled laborers in Silicon Valley)

*“Universe of cases”* with a non-zero probability of selection into your study. (consider phone numbers as sampling frame – for all phone users versus all adult Americans versus all first generation Southeast Asians 65 and older.)

The sample is the **actual list** from which measurement is sought. In most cases, a subset of sampling frame (e.g., subset of phone numbers, email addresses, postal addresses).

The **successfully measured** cases in your sample. Non-respondents is the complement in the sample.

**Post-measurement weights** to fit on population parameters. (usually benchmarked to demographic characteristics, using Census data).
Total Survey Error Approach

• **Measurement error**: observational gap btw. ideal measurement and observed response.

• **Coverage error**: non-observational gap btw. target population and sampling frame.

• **Sampling error**: non-observational gap btw. sampling frame and the sample.

• **Non-response error**: non-observational gap btw. sample and respondent pool.

• **Processing error**: observational gap btw. variable construction and observed response (including coding, data entry, transcription, disclosure avoidance errors).

• **Adjustment error**: non-observational errors from mistakes in assigning post-survey adjustment (e.g., variables used in weights).
Primers

→ Surveys and "public opinion" as a concept
→ Parallel processes and dual inferences in survey research
→ Sampling
Sampling and Coverage

UNDERCOVERAGE:
In the target population but missing from the frame (i.e.: non telephone household).

OVERCOVERAGE: In the frame but not members of the target population (i.e.: business telephone numbers)

FRAME POPULATION

Refused to Respond
Not Reachable
Not Capable of Responding

SAMPLED POPULATION
Sampling Methods

**PROBABILITY**
- random selection process
- generalizeability
- more expensive, and time consuming
- statistical analysis is more straightforward (e.g., known standard errors).
- amenable to hypothesis-testing

**NON-PROBABILITY**
- non-random selection process
- limited generalizeability
- less expensive, often easier and more convenient
- statistical analysis more complicated (e.g., unknown standard errors).
- more amenable to hypothesis-generation and mechanism testing
Hard to Sample Populations

• Racial minorities
• Immigrants
• Indigenous populations
• Sexual and gender minorities
• Linguistic and cultural minorities
• Institutionalized populations (hospitals, prisons, dorms, etc.)
• Mobile and migrant populations (homeless and refugee)
• Populations affected by natural disasters
• Populations in zones of armed conflict
• Stigmatized populations
• Populations that distrusts authority and science
Primers

→ Surveys and "public opinion" as a concept
→ Parallel processes and dual inferences in survey research
→ Sampling
→ Measurement
Measurement and "Non-attitudes"

- **Converse (1964):** “The Nature of Belief Systems in Mass Publics.”
- **Data:** ANES 1956-58-60 Panel
  - **Finding:** only 2.5% showed an ideologically consistent point of view, across items and panel waves. Overall response instability and incoherence.
  - **Conclusion:** The liberal-conservative continuum is too abstract and beyond “the man in the street.” Thus there is no underlying belief structure for most people, just “non-attitudes.”
A Model of Survey Response

Comprehension: Interpret the question
Retrieval: Search for relevant information
Judgment: Integrate information and make estimate
Response: Map judgment onto response category
Problems in Answering

- Failure to encode the information sought
- Misinterpretation of questions
- Forgetting and other memory problems
- Flawed judgment or estimation strategies
- Problems in formatting an answer
- More or less deliberate misreporting
- Failure to follow instructions
Measurement and Satisficing

• Survey respondents as satisficers, not optimizers
• Forms of weak satisficing:
  • Primacy and recency effects
  • Acquiescence bias
• Forms of strong satisficing:
  • Status quo endorsement, midpoint in rating scales, "straight-lining," "Don’t know" responses, “mental coin-flipping”
• Precipitants:
  • task difficulty, respondent ability, respondent motivation
General Principles of Measurement

• Do not reinvent the questionnaire wheel.
• Context when possible: topic, definitions, time frame, specific task (e.g., “select just one”)
• Avoid complex concepts or words or define the complexity
  • “tired” / “exhausted,” “work” / “employment”
  • “people who live here” vs. “occupants in this household”
• Other things to avoid:
  • Shorthand (e.g., abbreviations)
  • Negative wording (e.g. “how often do you not vote?”)
  • Double negatives (e.g., “do you agree or disagree that Obama should never not use the term ‘climate change’?”)
• Avoid the double-barrel ("do you want to be rich and famous?")
Measurement Pitfalls

• Respondents won’t always have answers:
  • “How much is your house currently worth?”

• Respondents may not always think about the construct the way you ask about it:
  • How many calories a day do you consume?
  • How many miles from your home is the nearest hospital?

• Respondents may not know about others:
  • “How many of your neighbors oppose the new park?”
  • “Does your mother enjoy the activities in her nursing home?”

• Respondents may not be able to recall:
  • How many different types of participation did you engage in last year?
To Learn More ...
Polling and Immigration Survey Primer

Sampling: Polling AAPIs
Evolution of Surveys of AAPIs

1984 IGS CA poll
- Asian + Latino, surname list + Korean oversample (only English + Spanish)

1990s LA Times polls
- LA / Orange CO, Chinese, Filipino, Korean, Japanese, Vietnamese (Asian languages)

2000-1 PNAAPS
- 5 MSA sample (NYC, Chicago, Honolulu, LA, SF), dual frame (targeted zip RDD + list)

2008 NAAS
- Nationally representative
- 5k+ completes 6 primary gps
- 8 interview languages

2012 NAAS
- 6k+ completes
- 11 languages
- SE Asian, NHPI Wh, Bk, Ltnx
- Community partnership

2016-7 NAAS
- Pre: voting and engagement
- Post: attitudes, experiences
- So. Asian, Wh, Bk, Latinx ++
- NSF
2008 NAAS

- **N = 5,159**
- **Mode = telephone (landline)**
- **Field dates = 8/18 to 10/29, 2008**
- **Sample = national sample and regional (CA, NJ/NY, “new destinations”)**
- **Languages:** English + Vietnamese, Korean, Mandarin, Cantonese, Tagalog, Japanese, Hindi (40% in non-English language)
- **Sample by groups:** 1,350 Chinese, 1,150 Asian Indian, 719 Vietnamese, 614 Korean, 603 Filipino, 541 Japanese, and 182 “Other Asian.”
- **Sampling frame:** list (with nominal RDD for comparison)
2012 NAAS  + Community Partner

→  **N = 6,257**

→  **mode = telephone** (82% landline, 12% cell, 6% VOIP)

→  **Field dates =** July 31 to October 20, 2012.

→  **Languages:** English, Vietnamese, Korean, Cantonese, Mandarin, Hmong, Khmer, Japanese, Tagalog, Thai, Hindi, and Spanish.

→  **Sample:** national, with oversamples of SE-Asians, NHPIs and comparison samples of whites, blacks, Latinos.

→  **Sample by sub-groups:** 827 Asian Indians, 743 Chinese, 633 Koreans, 599 Filipinos, 537 Vietnamese, 525 Japanese, 319 Hmong, 305 Cambodians, 251 other Asians, 419 Native Hawaiians, 152 other Pacific Islanders, 350 Whites, 309 African Americans, 308 Latinos
2016 Pre-Election NAAS

- **N** = 3,882 (2,238 AAPI)
- **Mode:** telephone (72% landline, 28% cell)
- **Field dates:** August 10 to September 29, 2016
- **Languages:** English, Vietnamese, Korean, Cantonese, Mandarin, Hmong, Khmer, Tagalog, Japanese, Laotian, Hindi, Spanish
- **Sample:** National with oversamples of SE-Asians + comparison groups
- **Sub-groups:** Cambodian (149), Chinese (352), Filipino (252), Hmong (325), Indian (307), Japanese (175), Korean (336), Vietnamese (342)
- **Comparison groups:** NHPIs (305), Whites (456), African Americans (392), Latinos (410), mixed race (54)
2016-7 Post-Election NAAS

• N = 6,448 (4,393 AAPI)

• Mode: telephone (63% landline, 37% cell)

• Field dates: Nov. 10, 2016 to Mar. 2, 2017

• Languages: English, Vietnamese, Korean, Cantonese, Mandarin, Khmer, Hmong, Japanese, Tagalog, Hindi, Urdu, Spanish

• Sample: National with oversamples of SE-Asians and So Asians + comparands

• Sub-groups: Bangladeshi (320), Cambodian (401), Chinese (475), Filipino (505), Hmong (351), Indian (504), Japanese (517), Korean (499), Pakistani (320), Vietnamese (501)

• Comparison groups: Latino (1,126), Black (401), White (408), NHPI (120)
Contribution of the NAAS

• Sampling:
  • Who: national + regional; "Big Six" initially; other key subgroups (SE Asians, So Asians, NHPIs); comparison groups.
  • How: list, mixed cell, language diversity, culturally competent firm.

• Measurement:
  • Replication of core ANES (then GSS) items
  • Within-group measures
  • Measurement on AAPI-specific issues
  • Comparison across groups and over time
## Example: 2016 Exit Polling

<table>
<thead>
<tr>
<th>race</th>
<th>clinton</th>
<th>trump</th>
<th>other/no answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>white 71%</td>
<td>37%</td>
<td>57%</td>
<td>6%</td>
</tr>
<tr>
<td>black 12%</td>
<td>89%</td>
<td>8%</td>
<td>3%</td>
</tr>
<tr>
<td>latino 11%</td>
<td>66%</td>
<td>28%</td>
<td>6%</td>
</tr>
<tr>
<td>asian 4%</td>
<td>65%</td>
<td>27%</td>
<td>8%</td>
</tr>
<tr>
<td>other race 3%</td>
<td>56%</td>
<td>36%</td>
<td>8%</td>
</tr>
</tbody>
</table>

24558 respondents

Source: Edison exit polls
Did Trump Really Outperform Romney?

Source: NEP and VNS exit poll data, 1992-2016

% Democratic vote share

- Blacks
- Latinos
- Asian Ams
- Whites

31 43 42 41 43 39 37 83 84 90 88 95 93 88 61 72 62 58 67 73 71 66 65

Source: NEP and VNS exit poll data, 1992-2016
### 2016 Primary Vote (NAAS)

“Which presidential candidate did you vote for in your state’s primary or caucus?

Only 18% reported voting for Trump; 76% for Clinton or Sanders

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Clinton</th>
<th>Sanders</th>
<th>Trump</th>
<th>All others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>57%</td>
<td>19%</td>
<td>18%</td>
<td>6%</td>
</tr>
<tr>
<td>Asian Indian</td>
<td>69%</td>
<td>18%</td>
<td>7%</td>
<td>5%</td>
</tr>
<tr>
<td>Chinese</td>
<td>61%</td>
<td>22%</td>
<td>15%</td>
<td>3%</td>
</tr>
<tr>
<td>Filipino</td>
<td>50%</td>
<td>16%</td>
<td>22%</td>
<td>13%</td>
</tr>
<tr>
<td>Japanese</td>
<td>55%</td>
<td>17%</td>
<td>23%</td>
<td>5%</td>
</tr>
<tr>
<td>Korean</td>
<td>73%</td>
<td>13%</td>
<td>14%</td>
<td>1%</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>55%</td>
<td>17%</td>
<td>24%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Source: Fall 2016 National Asian American Survey
“If the election were being held today would you be inclined to vote for Hillary Clinton, Donald Trump, or some other candidate?”

AAPI voters favored Clinton over Trump by a 4-to-1 margin

Source: Fall 2016 National Asian American Survey
2016 Vote Recall (NAAS)

- "Thinking about the past November election for President, did you vote for Hillary Clinton, Donald Trump, or some other candidate?
- AAPIs reported voting Clinton over Trump by a nearly 3-to-1 margin.

<table>
<thead>
<tr>
<th>Group</th>
<th>Hillary Clinton</th>
<th>Other Candidate</th>
<th>Donald Trump</th>
<th>DK/Refused</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>61%</td>
<td>5%</td>
<td>22%</td>
<td>12%</td>
</tr>
<tr>
<td>Asian Indian</td>
<td>66%</td>
<td>4%</td>
<td>14%</td>
<td>17%</td>
</tr>
<tr>
<td>Cambodian</td>
<td>64%</td>
<td>11%</td>
<td>14%</td>
<td>11%</td>
</tr>
<tr>
<td>Chinese</td>
<td>55%</td>
<td>4%</td>
<td>28%</td>
<td>13%</td>
</tr>
<tr>
<td>Filipino</td>
<td>58%</td>
<td>6%</td>
<td>26%</td>
<td>9%</td>
</tr>
<tr>
<td>Hmong</td>
<td>67%</td>
<td>4%</td>
<td>10%</td>
<td>19%</td>
</tr>
<tr>
<td>Japanese</td>
<td>64%</td>
<td>7%</td>
<td>19%</td>
<td>10%</td>
</tr>
<tr>
<td>Korean</td>
<td>71%</td>
<td>6%</td>
<td>17%</td>
<td>6%</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>53%</td>
<td>5%</td>
<td>30%</td>
<td>12%</td>
</tr>
</tbody>
</table>

Source: 2017 National Asian American Survey
2016 Election Eve Poll

2,391 Asian American voters

National sample of 863 (± 2%)

State samples in CA, FL, IL, NV, NC, PA, TX, VA (± 6.2 to 6.8%)

Ethnic targets of Chinese, Indian, Japanese, Korean, Filipino, Vietnamese (± 4.4 to 6.7%)

Interviews in English, Chinese, Filipino, Korean, Vietnamese

Field dates 11/1-11/7/16
More Credible Numbers

→ Asian languages at the point of contact. Landline, cellphone, and online.

→ Sample on extreme high-propensity vote history plus new registrants;

→ Sample with dedicated lists for high-propensity to be Asian (geography, surnames, consumption, etc.)

→ Screen on completed vote or certainty to vote;

→ Past samples validated at 88% positive for cast ballot.
Asian Americans Are Turning Right

GOP works to win over Asian Americans — and draws ‘race-baiting’ charges

Still solidly blue. But less so than before.
Keys to Surveying AAPIs: Sampling

• RDD is cost-prohibitive (nationally); telephone also $$$ cf. online

• Coverage issues with list samples:
  • Propensity-based on names and clustering based on geography works for some subgroups, not all.

• Coverage issues with language: 3 in 4 adults are foreign-born, 1 in 3 are Limited English Proficient
  • But increasingly hard to contact and get cooperation from LEP AAPIs.

• Weights are tricky (complex design effects, post-stratification due to non-response + due to vendor list coverage, response rate bias, etc.)

• Unknown if known sources of response rate bias apply to AAPIs.
Keys to Surveying AAPIs: Measurement

→ More accurate measures of key indicators (vote choice, mobilization).

→ Better measures of what is actually salient and relevant (e.g., issue agenda, candidate choice, sources of mobilization, knowledge).

→ Context-specific measures re AAPI experiences and attitudes (discrimination, affirmative action).

→ Within-group (AAPI sub-groups), between-group (AAPIs to others), and over-time comparisons.
Polling and Immigration

Sampling: Polling AAPIs

Sampling: Wild Wild West of Online Polling

Survey Primer
An Existential Threat to Survey Research?

Response rate by year (%)

Response rates to surveys, %

Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.


Source: Pew Research Center

Source: The Economist
Distinguishing Sources of Non-Response

- **Contact rate**: # someone in household reached ÷ eligible units. (3 AAPOR definitions)
- **Refusal rate**: # of refusals or break-offs ÷ eligible units. (3 AAPOR definitions)
- **Cooperation rate**: # all cases interviewed ÷ eligible units contacted. (4 AAPOR definitions)
- **Response rate**: # completed ÷ eligible units
  - 6 AAPOR definitions, varying with treatment of partial interviews and cases of unknown eligibility
Surveys Face Growing Difficulty Reaching, Persuading Potential Respondents

<table>
<thead>
<tr>
<th>Year</th>
<th>Contact rate (%)</th>
<th>Cooperation rate (%)</th>
<th>Response rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>90</td>
<td>43</td>
<td>36</td>
</tr>
<tr>
<td>2000</td>
<td>77</td>
<td>40</td>
<td>28</td>
</tr>
<tr>
<td>2003</td>
<td>79</td>
<td>34</td>
<td>25</td>
</tr>
<tr>
<td>2006</td>
<td>73</td>
<td>31</td>
<td>21</td>
</tr>
<tr>
<td>2009</td>
<td>72</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>2012</td>
<td>62</td>
<td>14</td>
<td>9</td>
</tr>
</tbody>
</table>

Contact rate (percent of households in which an adult was reached)
Cooperation rate (percent of households contacted that yielded an interview)
Response rate (percent of households sampled that yielded an interview)

Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.

PEW RESEARCH CENTER 2012 Methodology Study. Rates computed according to American Association for Public Opinion Research (AAPOR) standard definitions for CON2, COOP3 and RR3. Rates are typical for surveys conducted in each year.
## High-Effort Surveys Increase Contact and Cooperation Levels

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
<th>2003</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Landline</td>
<td>Landline</td>
<td>Landline</td>
</tr>
<tr>
<td><strong>Contact rate</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(% of households in which an adult was reached)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard survey</td>
<td>90</td>
<td>79</td>
<td>62</td>
</tr>
<tr>
<td>High-effort survey</td>
<td>94</td>
<td>91</td>
<td>86</td>
</tr>
<tr>
<td><strong>Cooperation rate</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(% of households contacted that yielded an interview)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard survey</td>
<td>43</td>
<td>34</td>
<td>16</td>
</tr>
<tr>
<td>High-effort survey</td>
<td>72</td>
<td>58</td>
<td>32</td>
</tr>
<tr>
<td><strong>Response rate</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(% of households sampled that yielded an interview)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard survey</td>
<td>36</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>High-effort survey</td>
<td>61</td>
<td>50</td>
<td>27</td>
</tr>
</tbody>
</table>

**PEW RESEARCH CENTER 2012 Methodology Study.** Rates computed according to American Association for Public Opinion Research (AAPOR) standard definitions for CON2, COOP3 and RR3.
Factors Affecting Cooperation

- People are too busy, too hunkered down, too self-absorbed, too wary of scams and unwanted intrusions, too mindful of privacy to cooperate.

- Cooperation is affected by:
  - Level of effort used in recruiting respondents
  - Mode of data collection and interviewer skill (if interviewer-administered)
  - Incentives to participate and assurances re purpose and privacy
  - Survey content: length, sensitivity, cognitive load.
  - Sample characteristics: e.g., homeless, undocumented, Trumpers, 1%ers
  - Respondents' interest in the topic of the survey
Cooperation in Phone Surveys

2020 Pew Study finds that only 19% of Americans pick up when an unknown number calls their cellphone.

Men, non-whites, young adults, and lower income Americans are more likely to pick up the call.

Most people check for a voicemail, but 1 in 7 ignore the call altogether.
Representativeness in Phone Surveys

Less educated, Latinx / AAPIs, young adults typically underrepresented in phone surveys.
Keeter et al (2000) compared two surveys with same questionnaire but different field procedures and different response rates.

<table>
<thead>
<tr>
<th>Design Feature</th>
<th>Standard</th>
<th>Rigorous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of field period</td>
<td>5 days</td>
<td>8 weeks</td>
</tr>
<tr>
<td>Respondent rule</td>
<td>Ask for youngest male at home; if no male at home, oldest female</td>
<td>Random selection</td>
</tr>
<tr>
<td>Advance letter</td>
<td>None</td>
<td>Yes, with $2 bill</td>
</tr>
<tr>
<td>Interviewers</td>
<td>Less experience More minority More female</td>
<td>More experience More white More male</td>
</tr>
</tbody>
</table>
Much Ado About Nothing?

• "standard" contact rate = 69%; cooperation rate = 58%; response rate = 36%
• "rigorous" contact rate = 92%; cooperation rate = 74%; response rate = 61%
• Results: Significant differences in 14/91 cases; mean difference (all 91 items) = 2%; Largest difference (9 percent) = interviewer rating of respondent interest
• Weakness: confounds many variables (respondent rule, advance letter, race and experience of interviewers, etc.)
• But is a 2000 study with 36% as a "low" response rate valid in a new regime of 1-5% response rates and online surveys with no calculable response rate?  
  *What do more recent data have to say?*
No Clear Partisan Bias

Surveys like the GSS, with high response rates (50-60%) track pretty to Pew Research polls with low response rates on political measures like partisanship.

Source: Pew Research Center (2017)
Similar re Ideology, Religion

% of respondents describing their political views as:

Pew Research Center Moderate
Pew Research Center Conservative
GSS Moderate
GSS Conservative

Pew Research Center Liberal


% of respondents describing their religious affiliation as:

Protestant Catholic Unaffiliated

Pew Research Center surveys General Social Survey

Source: Pew Research Center (2017)
## Overall, Minor Differences re Benchmarks

<table>
<thead>
<tr>
<th>% who say they ...</th>
<th>Benchmark</th>
<th>Pew Research Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have health insurance</td>
<td>87%</td>
<td>89%</td>
</tr>
<tr>
<td>Have a driver’s license</td>
<td>85%</td>
<td>86%</td>
</tr>
<tr>
<td>Are employed full or part time</td>
<td>62%</td>
<td>62%</td>
</tr>
<tr>
<td>Rate own health as excellent or very good</td>
<td>51%</td>
<td>50%</td>
</tr>
<tr>
<td>Are married</td>
<td>35%</td>
<td>35%</td>
</tr>
<tr>
<td>Have one or more children in household</td>
<td>20%</td>
<td>23%</td>
</tr>
<tr>
<td>Have a household income more than $50,000</td>
<td>18%</td>
<td>20%</td>
</tr>
<tr>
<td>Live in a one-adult household</td>
<td>16%</td>
<td>17%</td>
</tr>
<tr>
<td>Have a household income less than $20,000</td>
<td>11%</td>
<td>15%</td>
</tr>
<tr>
<td>Received food stamps or used a food stamp benefit last year</td>
<td>14%</td>
<td>14%</td>
</tr>
<tr>
<td>Have lived in current house for less than one year</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td>Smoke cigarettes every day</td>
<td>2%</td>
<td>6%</td>
</tr>
<tr>
<td>Are not a United States citizen</td>
<td>9%</td>
<td>13%</td>
</tr>
<tr>
<td>Received state or fed. unemployment compensation in last year</td>
<td>0%</td>
<td>6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% of U.S. adults who say they ...</th>
<th>Benchmark</th>
<th>Pew Research Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are a U.S. citizen</td>
<td>92%</td>
<td>92%</td>
</tr>
<tr>
<td>Have health insurance</td>
<td>88%</td>
<td>91%</td>
</tr>
<tr>
<td>Lived at same address one year ago</td>
<td>77%</td>
<td>78%</td>
</tr>
<tr>
<td>Only speak English at home</td>
<td>66%</td>
<td>68%</td>
</tr>
<tr>
<td>Voted in 2020 (among citizens)</td>
<td>66%</td>
<td>67%</td>
</tr>
<tr>
<td>Had at least one COVID-19 vaccine shot</td>
<td>66%</td>
<td>67%</td>
</tr>
<tr>
<td>Own their home</td>
<td>67%</td>
<td>68%</td>
</tr>
<tr>
<td>Live in a one-family, detached house</td>
<td>56%</td>
<td>58%</td>
</tr>
<tr>
<td>Worked for pay last week</td>
<td>52%</td>
<td>53%</td>
</tr>
<tr>
<td>Have had high blood pressure</td>
<td>33%</td>
<td>35%</td>
</tr>
<tr>
<td>Live in household that has at least one child</td>
<td>32%</td>
<td>33%</td>
</tr>
<tr>
<td>Received Social Security last year</td>
<td>22%</td>
<td>28%</td>
</tr>
<tr>
<td>Are a parent of a child in their household</td>
<td>26%</td>
<td>28%</td>
</tr>
<tr>
<td>Live in a single-adult household</td>
<td>15%</td>
<td>18%</td>
</tr>
<tr>
<td>Have a food allergy</td>
<td>9%</td>
<td>13%</td>
</tr>
<tr>
<td>Currently or formerly serve/served in military</td>
<td>9%</td>
<td>12%</td>
</tr>
<tr>
<td>Belong to a labor union</td>
<td>6%</td>
<td>10%</td>
</tr>
<tr>
<td>Smoke cigarettes every day</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Vape every day</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>Were unable to work because of COVID-19</td>
<td>1%</td>
<td>3%</td>
</tr>
</tbody>
</table>


"What Low Response Rates Mean for Telephone Surveys"

Pew Research Center 2015

Note: A total of 10,606 panelists responded out of 11,899 who were sampled (91%). However, the cumulative response rate accounting for attrition and nonresponse to panel recruitments is 3%.

Source: Survey of U.S. adults conducted June 14–27, 2021. See the "Benchmark sources” linked appendix for details on benchmark figures.

Pew Research Center 2021
RR and Political Engagement

% of respondents who are registered to vote ...

<table>
<thead>
<tr>
<th>Year</th>
<th>Pew Research Center</th>
<th>CPS Voting Supplement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>80%</td>
<td>70%</td>
</tr>
<tr>
<td>2000</td>
<td>73%</td>
<td>63%</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012 ’14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

% who say they ...

- Are registered to vote
- Always vote in local elections
- Contacted or visited a public government official in the last year

Note: Voting registration from the CPS was modeled using state-level voting data and the Hur-Achen adjustment. See methodology for details.

Source: Voter registration is an aggregate measure drawn from Pew Research Center telephone surveys in September and October 2014 and the Current Population Voting and Registration supplement. Pew Research Center estimates for contacting a public official and voting in local elections come from a survey conducted August 23, 2015 - September 2, 2016. The benchmark estimates come from the 2013 CPS Civic Engagement supplement.

“What Low Response Rates Mean for Telephone Surveys”

Source: Pew Research Center (2017)
RR and Civic Engagement

% who say they ...

- Trust all or most of the people in neighborhood
- Have done volunteer activities through or for an org, in the past year
- Talk with neighbors basically every day or a few times a week
- Have worked with people in neighborhood to fix a problem in the past year
- Have participated in school group or community org, during past year
- Have participated in sports or recreation org, during past year
- Have participated in a service or civic organization during the past year

• Biggest differences between high RR benchmark surveys and Pew surveys found for civic engagement.

• *Is survey response akin to a measure of civic engagement?*

Source: Pew Research Center (2017)
"Evaluating Online Nonprobability Surveys" (Pew, 2016)

- 10 nonprobability survey vendors estimates compared to federal population benchmarks: bias ranges from 5.8% to 10.1%

- Best performing vendor conditioned their sample on political variables (party, ideology, interest, registration). All sample overestimated volunteering (by 13 to 33%).

- Biases from weighted survey estimates compared to federal benchmarks were especially large for Latinos (8.3 to 19.8%) and also larger for blacks (8 to 14%). Biases also large for younger adults (7.1 to 16% for 18-29 y.o.)

- In multivariate regressions, "marginal effects associated with race and ethnicity are rarely correct."
The old conventional wisdom: telephone surveys are better than online surveys for parameter estimates of target populations of interest.

**Today's new frontier:** Which is better? Which is worse? A probability-based telephone sample with a 1% response rate or non-probability based online sample?

The value of both depend very heavily on adjustment weights. And do you weight on outcomes of interest, like civic and political engagement?
How Public Polling Has Changed in the 21st Century
Number of national public pollsters in the U.S. using method(s)

By 2022 the number of active pollsters more than doubled and methods diversified

Other
Online opt-in, probability-based panel and live phone
Online opt-in and live phone

Online opt-in only

IVR alone or with other methods
ABS (USPS) with multiple modes alone or with other methods
Probability-based panel and live phone
Probability-based panel only
Live phone (RBS) only
Live phone (RDD) only

About 29 national public pollsters were active in 2000, and nearly all used live phone

Note: Figures represent the number of active national public pollsters in each year and the method(s) that they used. IVR refers to interactive voice response, also known as robo-polling. ABS refers to address-based sampling. RBS refers to voter registration-based sampling. RDD refers to random-digit-dial sampling.

Source: Pew Research Center analysis of external data. See Methodology for details.

“How Public Polling Has Changed in the 21st Century”
% of pollsters who changed how they sample or interview people in national public polls in the U.S. during the two-year interval

Note: In this study change refers to using a different sample source or a different mode of interviewing. Figures for each interval are based on the set of pollsters that released at least one national public poll in both the starting year and ending year of the interval. Source: Pew Research Center analysis of external data. See Methodology for details. “How Public Polling Has Changed in the 21st Century”

PEW RESEARCH CENTER

% of national public pollsters in the U.S. using this many methods in polls they released each year

Note: A pollster is coded as using more than one method if they used more than one type of sample source (e.g., registered voter file, random-digit dial) or more than one interview mode (e.g., online, live phone). Source: Pew Research Center analysis of external data. See Methodology for details. “How Public Polling Has Changed in the 21st Century”

PEW RESEARCH CENTER
Do you know where your survey data come from?

Outsourcing data collection poses huge risks for public opinion

by Peter K. Enns and Jake Rothschild
from the Cooperative Election Study (YouGov)

The sample drawn for the CCES were chosen from the YouGov Panel, along with the Dynata, Critical Mix, and Prodege panels using a six-way cross-classification (age × gender × race × education × region × sample source). All respondents who completed the pre-election

from the AP VoteCast (NORC)

Nonprobability Sample

Nonprobability participants will include panelists from Dynata or Lucid, including members of its third-party panels. In addition, some registered voters will be selected from the voter file, matched to email addresses by V12, and recruited via an email invitation to the survey. Digital fingerprint software and panel-level ID validation is used to prevent respondents from completing the AP VoteCast survey multiple times.

Polling data you can trust

Get accurate insight into who voted and why as soon as polls close on Election Day with data from AP VoteCast. For nearly 175 years, AP has tabulated election results and declared winners in U.S. elections. Now our rich and robust voter survey helps tell the whole story of the American democracy.
The Data Outsourcing Problem

"Many of the most prominent companies in the industry were using Lucid to get data. Lucid, in turn, gets data for these companies by reaching out to hundreds of different data providers — potentially unknown to the original client."

Lucid partners with more than 350 global suppliers.

- DISQO
- ondevice research
- tap research
- prodege
- amee
- GMO research
- BOVITZ
- branded research inc.
- Dalia
- maru/
- Tapjoy
- dataSpring
Enns and Rothschild Recs: Ask ...

→ Are the data / respondents outsourced, or collected directly?
  → If "yes," request a full list of potential respondent sources, then run quality checks (e.g., AAPOR Transparency Initiative, Roper Transparency Score)
  → Also ask if any of those respondent sources collect directly or further outsource for respondents.
→ Do any sources route respondents to complete consecutive surveys? (re: survey fatigue and satisficing)
→ How many surveys can respondents take each week?
→ What happens to someone who does not qualify for a survey? (incentives to falsify qualifications, survey fatigue)
→ What is the compensation for participation, down the line of outsourcing?
Thank you!

taelee@fas.harvard.edu